

Goal-UML Approach For Modeling Data Warehouse Requirements Of Malaysian Rural Health Care

Sajaratul Huda Shafie, Azizah Ahmad, and Azman Ta'a

Universiti Utara Malaysia, Malaysia, amanhudashafie86@gmail.com, {azie, azman} @uum.edu.my

ABSTRACT

Malaysian health care authorities in providing information system for rural communities required to meet up with the challenges, to provide informed decision-making by utilizes their enormous data. However, only a few of these initiatives have their success stories as many are still struggling to justify the investments such as Data Warehouse System (DWS). Among the most reason of the failures were that DWS requirements were overlooked, leading to poor DWS deployments. This paper attempts to model DWS requirements using goal-oriented and UML-based approach. DWS requirements for Rural health care were modeled two-folds: decision making requirements that centered on stakeholders; and focused on organizational and decisional aspects. The model can guide DWS development in the process and data needed in rural health care strategic decision-making. This provides new insights and facilitates the improvement of new health care knowledge.

Keywords: Data Warehouse, Requirement Analysis, Goal-oriented, Business Intelligence, UML

I INTRODUCTION

The Malaysian health care is facing a lot of challenges as the rural dwellers are accessing to poor health care services (Ali, Dalpiaz & Giorgini, 2010). The consumers of rural health care in Malaysia are mainly those whose purchasing powers are very limited and they are very much dependent of free public services provided by the government. They also have limited choice of accessing other health care providers such as private hospitals, which provide better products and services. Another prominent issue is the present number of health care professionals in the rural health care organizations. On average the rural folks possess less doctors and nurses as compared to their urban and wealthier friends and leaving them in sicker and poorer conditions (Alencar et al., 2000).

Realizing these issues and to overcome these challenges, the Malaysian government has since put in place many initiatives including implementing information systems. These systems require them to gather health care information such as patients, diagnosis, treatment and others, have been storing a massive amount of information in databases throughout the country over the years. However, so

far proper programs are still not in place to analyze these so called 'Big Data' owned by them. Owning massive amount of data, a forward-thinking health care organizations must realize that data and, thus, a concept known as Data Warehouse System (DWS) or Business Intelligence (BI) is at the center of informed and precise decision-making that will improve their services (Bresciani et al., 2004). To achieve the full benefits of DWS in any organizations, there must be a strategic approach to defining decision-making requirements for DWS. The proper modeling of data, analytics and decision-making process should be done for the people in organizations take advantage of their data.

This paper identifies, analyzes decision-making requirements for rural health care center and finally, constructs a DWS requirement model for rural health care centers in Malaysia.

II DATA WAREHOUSE SYSTEM

DWS has emerged as an important area of study for both practitioners and researchers, reflecting the magnitude and impact of data related problems to be solved in contemporary business organizations including health care organizations (Nguyen, Perini & Tonella, 2007). Scholars from different school of thoughts have their own definitions of DWS, but generally it is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance (Ellis-Braithwaite et al., 2013). From a technology point of view, DWS is used for gathering, storing, analyzing, and providing access to data for better business decisions (Kumar & Singh, 2010). Some earlier scholars view DWS as a process of turning data into information and then into knowledge that can be used for good decision-making (Mazon, Pardillo & Trujillo, 2007).

With competitive and complex business environment today, organizations, private or public are under tremendous pressure. They need to respond quickly to changing conditions and be innovative in order to stay competitive. It requires them to be agile and make frequent strategic, tactical, and operational decisions. Making such decision may require considerable amounts of timely and accurate knowledge. Processing these in the framework of the required decisions needs quick, frequent and some computerized support, which is traced to DWS

(Giorgini, Rizzi & Garzetti, 2008). Making DWS a success story is the real challenge for organizations, among the main factors that mentioned in the literature is that DWS doesn't fulfill their decision-making requirements. Many DWS systems failed to deliver the promised 'knowledge' needed by decision makers.

Requirement plays a very important aspect of DWS implementation, where poor requirements management can lead to project failure. DWS project managers need to acquire necessary skill in order to optimize the results and to minimize problem during DWS requirements elicitations. DWS requirements analysis can be classified in two categories: i) data-driven approaches starting from analysis of data sources; user requirements impact on design by allowing the designers to select which chunks of data are relevant for decision making and by determining their structure according to the multidimensional model; and ii) requirement-driven start from determining the information requirements of end users, and how to map these requirements onto the available data sources. Data-driven approaches simplify the design, but they give end-user requirements a secondary role. On the other hand, requirement-driven approaches start by determining the information requirements of end users, which bring requirements to the foreground, but it requires larger efforts when designing as it needs a more structured method.

Information systems scholars have come up with techniques for modeling that help analyze and understand requirements for systems such as Context Diagram, Entity Relationship Diagrams and Sequence Diagram. But most of these modeling techniques only cater for the operational type of information systems, whereas DWS is a strategic information system that has different needs in terms of data as well as processes.

A. Requirement Modeling with Goal-oriented Approach

In the field of DWS design, Giorgini et al. (2003) proposed Agent-Goal-Decision-Information (AGDI) model to support the early and late requirements for development of DWSs from the stakeholders' perspective. This model also supports three interrelated modeling activities, namely, organization modeling, decision modeling and information modeling. Gupta et al. (2011) proposed a comprehensive method in order to determine information requirements of DWS users and match these requirements with the available data sources. The activity model represents the core component of a comprehensive methodology for information requirements analysis for DWS systems.

Requirement analysis is split in two which is early and late requirement analysis in designing phase (Horkoff & Yu, 2012). Maté, Trujillo & Franch (2014) stated that during the first phase, domain stakeholder and model will be identified as social actors, and in late requirement analysis, the conceptual model will be designed consisting of new actor and a number of dependencies with other actors. Lujan-Mora (2005) presented the comparison between different approaches to analysis the goal-oriented requirement models, to understand the ways in which procedural design choices affect results.

The sufficient requirements to support the decision making process are required to manage the DWS changes correctly. This will help the information provided will meet the business goals. However, the requirements for DWS are difficult to decide from the goal of the decision maker or organization since the current DWS modeling does not consider the goal concept in the model (Giorgini, Rizzi, & Garzetti, 2008). All the prerequisites must be squared up, identified and analyzed within early and late requirement analysis by using suitable modeling.

Requirement analysis is split in two which is early and late requirement analysis in designing phase (Wirtz et al., 2006). This stated that during the first phase, domain stakeholder and model will be identified as social actors, and in late requirement analysis, the conceptual model will be designed consisting of new actor and a number of dependencies with other actors, meanwhile in architectural and detail design phases focus on the DWS detailed design specification in the implementation phase.

In the perspective of goal-oriented approach, Giorgini, Rizzi & Garzetti (2008) presented the comparison between different approaches to analysis the goal-oriented requirement models, to understand the ways in which procedural design choices affect results. They were advocated goal-oriented to capture and link technical requirements to derive high-level or details user requirements using elicited goals, capture and compare alternative potential implementations. Several applications of goal-oriented techniques in different modeling frameworks, techniques, or methodology include KAOS, GBRAM, NFR, i*, Tropos, GRL, and AGORA.

Mazon, Pardillo & Trujillo (2007) proposed a GORE approach for modeling, organizational goals that the DWS supports and relating them to the information required and to use the i* modeling framework and the model driven architecture (MDA) in order to describe (i) how to model goals and information requirements for DWSs, and (ii) how to derive a conceptual multidimensional model that provides the required information to support the decision making process. Computation independent model (CIM) is specified by

using the i* modeling framework in order to model goals and information requirements for a DWS and then the conceptual multidimensional model of a DWS is derived from the CIM into the platform independent model (PIM).

B. Tropos Methodology

Tropos is an agent-oriented software engineering methodology (Giorgini, Rizzi & Garzetti, 2008). In DWS design, Tropos can be used as goal-oriented approach to requirement analysis in two perspectives of modeling which is organizational modeling and decisional modeling. Tropos has proposed a few steps in goal analysis named GRAnD as shown in Figure 1. GRAnD methodology that encompassed these two modeling supports from early phases of requirements analysis to detailed design that focused on the understanding of the environment where the DWS must operate, and provides communication between analyst and stakeholder in the decision-making process.

There are four phases in Tropos methodology consist of early requirements, late requirements, architectural design, and detailed design. Early requirement concerned about the intentions of stakeholder that underlie in DWS design, problem in the decision-making process on organizational setting, exploring system solution and alternatives, and it must be done before UML modeling.

Early requirement analysis proposes two main diagrams which are the actor diagram and the goal diagram in goal analysis. Tropos has adopted Eric Yu's i* model, which offer actors (agent, roles or position), goals, and actor dependencies in the early requirement modeling. Late requirement analysis defined the requirement specifications, including functional and non-functional requirements for the DWS design. Meanwhile, architectural design defined the interconnected through data, control and dependencies, and detailed design defined the behavior of each component.

Figure 1 shows the Tropos notation that used in DWS design. An actor is represented as stakeholder. In a health care domain, there are Ministry, Hospital and Clinic. An actor related by dependencies to the other actor to achieve the goal in actor diagram. Actor diagram is a model to show how the actor depending on another actor to respect the goal of the organization. Meanwhile, rationale diagram is used to represent the model goals and sub-goals of actors.

There are a few new concepts required in DWS context by using Tropos, which are facts, attributes, dimensions, and measures. Fact is a set of events that happen when the goal is achieved, attribute is a value provided when the fact is recorded to achieve the goals, dimension is a facts property to fulfill the

analysis goal and measures is a numerical property of a fact related to the decision making.

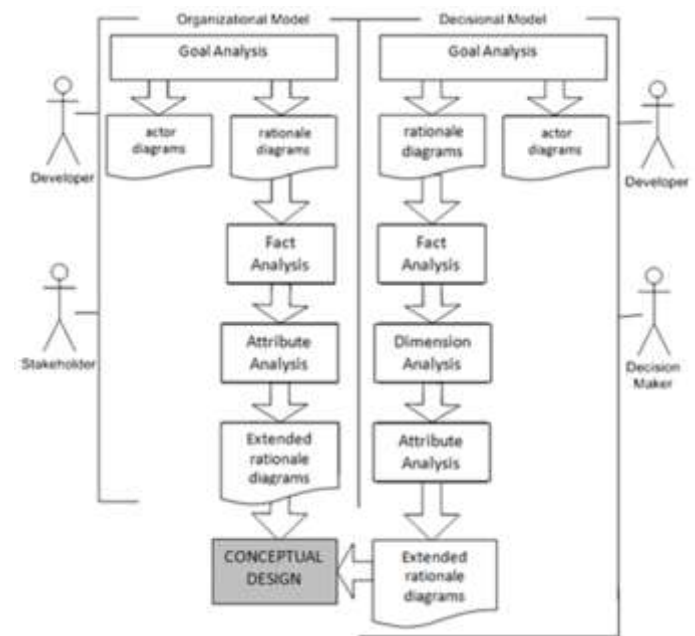


Figure 1. The GRAnD Approach

Figure 2 shows the notation that used in DWS design. An actor is represented as stakeholder. In a health care domain, there are Ministry, Hospital and Clinic. An actor related by dependencies to the other actor to achieve the goal in actor diagram. Actor diagram is a model to show how the actor depending on another actor to respect the goal of the organization. Meanwhile, rationale diagram is used to represent the model goals and sub-goals of actors.

There are a few new concepts required in DWS context by using Tropos, which are facts, attributes, dimensions, and measures. Fact is a set of events that happen when the goal is achieved, attribute is a value provided when the fact is recorded to achieve the goals, dimension is a facts property to fulfill the analytical goal and measures is a numerical property of a fact related to the decision making.

In the result sections, the phases are described in detail with reference to real case study, the Malaysian Rural Health Care Centers. The system used in hospital and clinic is not integrated with each other. This causes the management of information and decision-making between medical practitioners and staffs has a difficult time to organize and utilize these volumes of patient's data effectively and efficiently. This indicates it is a necessity for implementing the DWS concept in this institution's health care as a solution to integrate the system in order to collect and distribute the health care data.

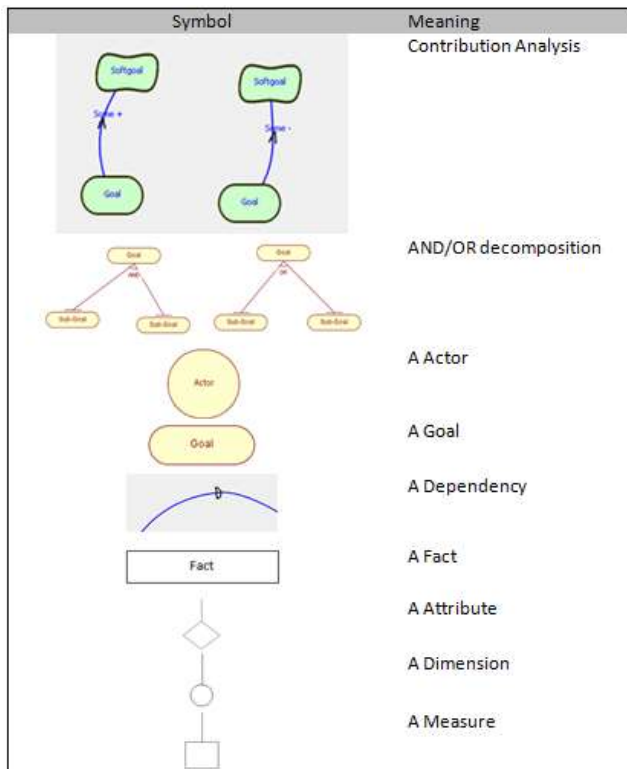


Figure 2. Notation for Actor and Rationale Diagram

III HEALTH CARE ORGANIZATIONS

The global health care environment has widely divergent perspectives on the use of data and information for decision-making. The ability to collect and analyze data garnered from the point of patient care has been impressive. Health Care delivery, however, has often been plagued by under funded, less advanced methods of collecting and analyzing data. Most providers continue to use (EHR) systems and strive to integrate systems that combine both clinical and administrative data. Through this transition, health care provider organizations can take advantage of this data and explore analytics as a competitive tool as a method to help provide better care, improved outcomes and safer, more effective decision making. Taken together, systems and data cannot solve all of the problems that face health care system alone. This requires an eye toward setting the strategy based on sound fundamentals along with policy decisions that govern the operations of health care environment. Hence, exploring and understanding the health care requirements of DWS is vital in order to achieve the goals.

There are a lot of issues in rural health care spanning from managing medical professionals related to clinical problems. The successful implementation of rural health care in Malaysia requires competence and professional benchmark, which embedded patient in the values and behavior. The rural health care staffs in

Malaysia have witnessed upgrading in knowledge and skills to sustain the right value and attitudes. One of the contemporary issues bordering Malaysian rural health care is the number of health workers and their distribution. The fact that the number has been increasing over the year, the effect is yet to be felt in catering for the rural dwellers while their demands are yet to be met. The issue of distribution of experience and skillful health care professionals in Malaysian rural areas can be tackled by applying DWS as a measure. Despite the great strides made in socioeconomic development, there are still remains pockets of disadvantaged communities such as aborigines in the remote rural areas. The disadvantaged and marginalized groups in the rural areas need to be alleviated in achieving successful rural health care services in Malaysia. Nevertheless requirement analysis for DWS health care systems, in this case rural health care, has not been given much attention so far. DWS requirement analysis is often neglected due to: i) the projects are long-term ones, in which most requirements cannot be stated from the beginning; and ii) requirements are poorly shared across organizations, unstable in time, and refer to information that must be derived from data sources. Several surveys indicated that a significant percentage of DWS projects fail because the requirement analysis is overlooked.

IV METHODOLOGY

This study employed a qualitative approach, where interview and documents, sampling methods were used to collect relevant health care data needed for the modeling process. The semi-structured interviews with 15 health care staffs including 8 medical doctors, 5 nurses and 2 administration staff from 5 rural health care centers in Kedah and Perlis were administered. The relevant documents were also sampled to get in depth views on health care data, processes and decision making activities. These interviews were analyzed by the content analysis method and the results were used to model DWS requirements. Goal-oriented requirements analysis and design (GRAnD) proposed by Giorgini, Rizzi & Garzetti (2008) were adapted to model the requirements and the modeling results are discussed in detail in the following sections.

V MODELING DATA WAREHOUSE

Based on content analysis result on interview data and document sampling, 2 types of modeling were performed. Based on GRAnD method, DWS health care were modeled according to organization in order

to get the stakeholder's point of view, and decisional to get their goal and decision-making activities.

A. Rural Health Care Organization Modeling

The organizational modeling represents the main data in the organization and comprise most relevant attribute to data sources. The organizational modeling consists of three types of analyses which is goal analysis, fact analysis, and attribute analysis and produced lists of facts and attributes.

Goal analysis

The first step in goal analysis represents the intentions of relevant stakeholder for the organization and their dependencies in actor diagram. Follow by analysis to decomposing the high level goal into sub-goal as shown in Figure 3.

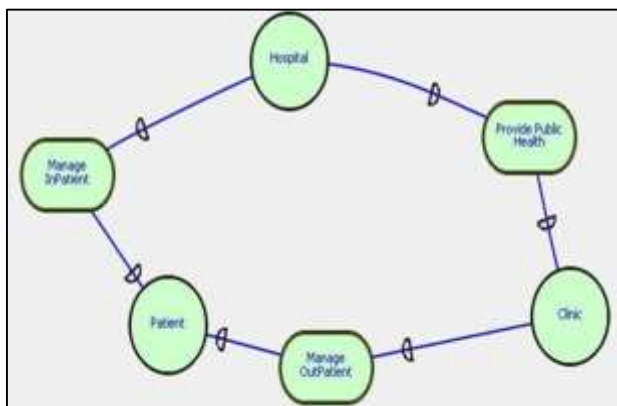


Figure 3. Actor Diagram for Rural Health Care.

The second step produced a rationale diagram for each actor in actor diagram for Rural Health Care as Figure 4. Goals are AND decomposed to sub goals and contribution link between goals are discovered. The analysis ends when all the goals of each actor have been analyzed and all the dependencies among actors are established in actor and rationale diagrams. The previous rationale diagram from goal analysis is used to identify all the relevant facts for the organization. At this phase of view, the analyst navigates the rationale diagram for each actor and extends it by associating goals with facts. The fact OPD Services is associated with the goal Provide OPD Services. The information is collected using two different types of template which are (Fact, Description) and (Goal, Fact).

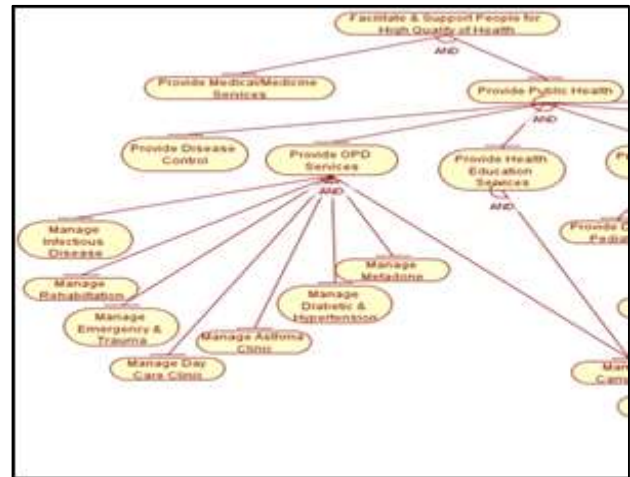


Figure 4. Rationale Diagram for Rural Health Care: Public Health (Organizational Modeling)

Attribute analysis

All the attributes that are given a value is identified when facts are recorded. Starting from the previous diagram, the entire related attribute for the sub-goals are explored. The attributes are simply data that associated with the goals. The information is collected by using table (Attribute, Goal, and Fact).

B. Rural Health Care Decisional Modeling

Decisional modeling consists of four types of analyses which is goal analysis, fact analysis, dimension analysis and measure analysis to support the decision making and produced lists of facts, dimensions and measures. Decisional modeling is focused on how the DWS supports the decisional process of the organization, and the requirements of the DWS from the perspectives of the decision maker. The previous diagram for organizational modeling is used to support the identification of the facts that to be associated with the decision maker goals.

Goal analysis

In this analysis phase, goal analysis starts with analyzing the actor diagram for the decision maker. A decision maker is identified and initial dependencies between them were established. The goals are the decomposed to produce rationale diagrams.

Fact analysis

A rationale diagram for organizational modeling is used to identify the facts and as associating it to goals of decision makers. Basically, the facts for rationale diagram are imported from a rationale diagram in organizational modeling.

Dimension analysis

In dimension analysis, each fact is related to the dimensions that decision maker considers to achieve the decisional goals. Dimensions are identified by

analyzing the goals and facts from a rationale diagram of decision maker.

Measure analysis

Finally, a set of measures is associated to each of facts from previous diagram. A measure is a numerical property of facts that relevant to the decision making process.

The organizational model produced by requirement analysis represents the main data that comprises the most relevant attributes that are part of the source database. On the other hand, the decisional model describes the decision maker's needs, which summarizing the role played by an actor associates to the facts, dimensions and measures. The requirements derived from organizational and decisional modeling are matched with the schema of source database to generate the conceptual schema for Malaysian Rural Health care. There are three phases involved, which are requirement mapping, hierarchy construction and refinement based on GRAnD. Facts, dimensions, and measures from extended rationale diagrams are mapped onto the source schema. Every dimension associated with a goal related and successfully mapped to the decisional model to the source is included in the conceptual schema. Figure 5 shows the finding of conceptual schema for rural health care DWS. The model shows that the dimensions are associated mapped to facts from decisional model to source schema with many-to-one relationship. Every measure associated mapped to a goal related to facts from a decisional model to the source schema and provide many-to-one relationship.

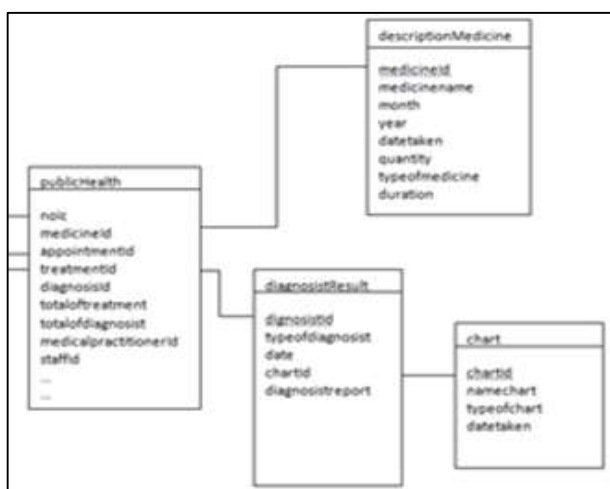


Figure 5. Conceptual Schema For Rural Health Care

VI CONCLUSION

This paper proposes a model of DWS requirements for Malaysian rural health care. A goal-oriented

methodology for requirement analysis in DWS is explored. The goal-oriented modeling is proposed to capture the requirements of DWS from stakeholders' perspective, which are organizational modeling and decisional modeling. The approach captured and analyzed the early requirements for the Malaysian health care focused on rural health centers by using the Tropos Methodology. The adoption of GRAnD in this modeling approach can help the designer to reduce the risk of project failure and at the same time, makes the design of DWS simpler. The modeling results proved that GRAnD can be used as a language as well as method to derive data warehouse conceptual schema.

REFERENCES

- Ali, R., Dalpiaz, F., & Giorgini, P. (2010). A goal-based framework for contextual requirements modeling and analysis. *Requirements Engineering*, 15(4), 439-458.
- Alencar, F. M., Castro, J., Cysneiros Filho, G. A., & Mylopoulos, J. (2000, July). From Early Requirements Modeled by the i* Technique to Later Requirements Modeled in Precise UML. In *WER* (pp. 92-108).
- Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., & Mylopoulos, J. (2004). Tropos: An Agent-Oriented Software Development Methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3), 203-236.
- Nguyen, D. C., Perini, A., & Tonella, P. (2007, May). A goal-oriented software testing methodology. In *International Workshop on Agent-Oriented Software Engineering* (pp. 58-72). Springer Berlin Heidelberg.
- Ellis-Braithwaite, R., Lock, R., Dawson, R., & Haque, B. (2013). Towards an approach for analysing the strategic alignment of software requirements using quantified goal graphs. *arXiv preprint arXiv:1307.2580*.
- Kumar, M., & Singh, Y. (2010). Stakeholders driven requirements engineering approach for data warehouse development. *Journal of information processing systems*, 6(3), 385-402.
- Mazon, J.-N., Pardillo, J., & Trujillo, J. (2007). A Model-Driven Goal-Oriented Requirement Engineering Approach for Data Warehouses. *LECTURE NOTES IN COMPUTER SCIENCE*, (4802), 255-264.
- Giorgini, P., Rizzi, S., & Garzetti, M. (2008). GRAnD: A Goal-Oriented Approach to Requirement Analysis in Data Warehouses. *Decision Support Systems*, 45, 4-21.
- Giorgini, P., Kolp, M., Mylopoulos, J., & Pistore, M. (2003). The Tropos methodology: An overview. *Methodologies And Software Engineering For Agent Systems*, Kluwer Academic Publishing (New York).
- Gupta, V., Chauhan, A., Kumar, A., & Taneja, S. (2011). UREM-A UML-Based Requirement Engineering Model for a Data Warehouse. In *Proceedings of the 5th National Conference*.
- Horkoff J. & E. Yu. (2012). Comparison and evaluation of goal-oriented satisfaction analysis techniques. doi: 10.1007/s00766-011-0143-y
- Maté, A., Trujillo, J., & Franch, X. (2014). Adding semantic modules to improve goal-oriented analysis of data warehouses using I-star. *Journal of Systems and Software*, 88, 102-111.
- Lujan-Mora, S. (2005). *Data Warehouse Design With UML*. Unpublished PhD, University of Alicante.
- Wirtz, K., Tauscher, M., Zwerenz, M., Munte, A., und Strategie, V., & Bayern, K. V. (2006). Data warehousing for Bavarian Out-Patient Public Health Care. In *ECEH* (pp. 263-274).